On the Conflation of Consciousness and Cognitive Complexity

Katrina Schleisman[1] and Michael Levin[2,3]
[1]Galois Inc.
[2]Tufts University
[3]Harvard University
katrina.schleisman@galois.com, michael.levin@tufts.edu

Abstract

The goal of detecting, measuring, and engineering machine consciousness depends on making explicit key metaphysical assumptions about the nature of consciousness that may be lying dormant in the minds of AI researchers. In this position paper, we lay out a small set of arguments that call into question a particular family of these assumptions. These assumptions all arise from the dominant paradigm of analytic materialist philosophy in combination with modern cognitive science: 1. consciousness is a property that emerges from complex matter, 2. consciousness depends on architectural cognitive complexity, and 3. consciousness can be measured as a function of behavioral cognitive complexity. Alternatives to these views and their implications for future machine consciousness research are discussed.

Introduction

The acceleration in AI capability and complexity over just the past few years has suddenly brought to the fore a question that was previously only of casual interest to most AI researchers, namely, *when will the machines wake up?* The current community of AI research now seeks to produce concrete, actionable, and scientific approaches to solve problems that have occupied the minds of philosophers throughout human history. Arguably, we would all like to be able to detect the appearance of machine consciousness if and when it arises, and to establish principled guidelines for how to relate to it once it's here. Some of us would like to purposefully engineer machine consciousness, while others would like to prevent, slow down, or set limits on the engineering of machine consciousness. Disagreements on this latter consideration are active, but it is all moot unless we can come to agreement on 1. basic definitions of consciousness, both informal and formal, and 2. basic methods and tools for measuring and detecting consciousness in machines.

Metaphysics and Matter

The first assumption we wish to address is the basic definition of consciousness in material terms. *Materialism* is the metaphysical stance that matter is the underlying reality of the universe. It dominates thinking in Western culture and is foundational to the history of science and the empirical focus of the scientific method (Schlegel-O'Brien 2024). Consequently, theories of consciousness that are based on materialism must be posed in terms of answers to a single question: *how does consciousness arise from matter?* Consciousness is often seen as a late comer, showing up (rarely) at some point in a material universe that was previously empty of mind and first-person experience. In our experience as scientists working in the 21st-century Western world, materialism is so pervasive that it is almost universally and implicitly accepted–despite the fact that science does not presume a specific metaphysical ontology. While it is not the aim of this paper to make any particular case against materialism, it is important to bring this metaphysical assumption into the light so that our discourse on the nature of consciousness and its manifestation in machines begins from known foundations.

To deny or to question materialism can take many forms, the most popular in Western analytic philosophy being *dualism*, which posits a reality in which mind and matter are fundamentally different yet co-mingling aspects of the universe. Consequently, theories of consciousness that are based on dualism do not face the same constraints as theories based on materialism–given the view that mind and matter are different in their essential nature, consciousness does not require an explanation in terms of matter. Instead, the key problems are understanding the nature of mind as distinct from matter and defining exactly how it is that two fundamentally different entities can successfully interact in a conscious embodied mind. Dualism is mainly familiar to people as "Cartesian dualism", eponym of 17th-century French philosopher Renee Descartes. However, there is a current resurgence of interest in exploring dualist models, in response to a need to understand the latent space of complex "regularities" and patterns that occur not only in mathematics, but in physics, biology, and cognitive science (e.g., the *Platonic Space Symposium*, Levin 2025a; Levin 2025b). Our purpose in raising the issue here is to invite AI researchers not to reflexively discount non-materialist frameworks and to develop their own arguments in support of or against them as they reason through how to best approach research in machine consciousness.

Going a step further in the metaphysical journey we come to *idealism*, in which the reality of the universe is proposed to be fundamentally subjective, conscious, or mental in nature. Consequently, theories of consciousness that are based on this view bear little resemblance to theories of consciousness developed through the materialist tradition, not least because they posit that matter does not, in fact, truly exist. This view is exemplified in Western analytic philosophy by 18th-century Irish philosopher George Berkeley, and much more broadly by the whole of Eastern philosophy via the foundational theory of Tantrism and further developed in Hinduism, Buddhism, Jainism, and other related schools of thought. The foundations of these Eastern philosophical traditions pre-date Western thought by several millennia, and as a result the body of written work passed down through history contains arguably the most sophisticated frameworks of consciousness available. The reason for this is that under a metaphysical worldview in which consciousness is the primary foundation of reality, consciousness then becomes the primary subject and object of relevance to study. This is completely in contrast to the Western analytical tradition, in which consciousness will be handled by the scientific method once we achieve complete understanding of scales of organization of matter.

That the materialist tradition is being stretched to its limits scientifically is perhaps most apparent in the field of quantum mechanics, where multiple experiments have demonstrated a puzzling dependence of observed phenomena on subjective perspective, i.e., on observation. This has led to the development of new, subjectivity-first theories in physics such as *relational quantum mechanics* (Rovelli 1996) and quantum Bayesianism or *QBism* (Fuchs 2014). The rejection of materialism has an even longer history in physics, with theoretical physicist Max Planck asserting: "I regard consciousness as fundamental. I regard matter as derivative from consciousness." (Planck 1931). Anti-materialist ideas are also influencing cognitive science theory, e.g., as *conscious realism*, which asserts the fundamental status of conscious experience and the fundamental role of conscious agents in creating, rather than representing or reflecting, what we commonly take to be objective reality (Hoffman 2019).

In all, this incomplete speed-run through metaphysics is intended solely to shake up the discussion of how to define and measure consciousness in machines. As we endeavor to make progress in unraveling the most fundamental mystery of our existence, we would all do well to question our most basic assumptions.

Architectural Cognitive Complexity

Another place where we find hidden assumptions is when we ask just *who* and *what* in the universe is conscious? We often start in agreement with the assertion that we, ourselves, are conscious, and then work our way through the ontology of things in the world until we reach a point where we begin to genuinely ponder whether that thing is conscious or not. This way of thinking has been handed down to us from Greek philosophers like Aristotle in *De Anima* (Shields 2016) progressing to the medieval concept of the Great Chain of Being (Lovejoy 1936). This ontology places humans near the top of a chain that descends from God, through animals, plants, and down to the bottom comprised of inorganic matter. The chain assumes that each level of the hierarchy subsumes those below it, such that we as humans possess all the qualities and capabilities of so-called lower-order animals, plants, and minerals, along with qualities those entities do not possess. What falls from this kind of ontology is the view that consciousness (or soul, as it was more commonly termed) only emerges at the highest levels of the chain. As a result, consciousness becomes, intentionally or unintentionally, associated with the complexity of an entity, and the complexity of an entity's architecture and function are taken as evidence for or against its consciousness. Rock? Definitely unconscious, we are prone to say. Human? Definitely conscious. Cockroach? More research needed. Especially because we don't have a good definition of complexity in the relevant sense. Mere incompressibility (a common definition of complexity) doesn't do the trick. Multi-scale heterarchies with specific causal structures (Watson, Levin, and Buckley 2022) are more likely, but the many assumptions and knowledge gaps behind "complexity" need to be kept in mind.

Because of our bias to consider humans and other entities near the top of the chain as uniquely conscious, or at least possessing a unique *level* of consciousness, we tend to search for clues that will reveal the relationship between consciousness and cognition by focusing on architectures and processes that we do not share with other, potentially unconscious, entities. Debates have waxed and waned over whether consciousness depends on the human brain's comparatively large neocortical sheet (Penfield 1938; Merker 2007), its large prefrontal cortex and associated functions of cognitive control and metacognition (Mathias 2022), cortical and sub-cortical structures that support natural language (Arbib 2001; Searle 2002), the hippocampus and its role in episodic memory (Behrendt 2013), and many others.

What we wish to claim here is that the evidence for a link between architectural or functional cognitive complexity and consciousness is tenuous. We might begin by considering whether we can establish credible evidence that consciousness varies across different types of entities and compare their relative cognitive complexity, but to do this we would need to make a priori assumptions about which entities are conscious and which are not. Instead, we restrict our analysis to just one such entity–humans. We proceed by first accepting the assertion that human beings are conscious (if you disagree with this, you can stop reading here) and then argue that humans display an enormous range of cognitive complexity. Despite this range, we do not observe a clear relationship between various degrees of architectural or functional cognitive complexity and the consciousness that supposedly depends on it.

Our first example is that of Helen Keller (Keller 2000). Due to illness, she lost her sight and hearing at the age of 19 months and could not speak. She famously learned to communicate through letters signed onto her palm by her teacher Ann Sullivan, who began working with Helen just prior to the age of 7. Helen Keller went on to become the first deaf-blind person to earn a college degree in the United States, at Radcliffe College in 1904. She went on to author her autobiography as a series of letters, *The Story of My Life*. What is remarkable in reading her writing is the degree to which she eloquently communicates her rich, conscious experiences in natural language despite lacking a direct connection with the sensory modalities that most of her readers possess. She writes about abstract concepts that are extremely difficult to imagine being transmitted as knowledge through palm signaling and describes objects in the world that she's never seen or heard. We believe that the case of Helen Keller illustrates that a mind with "less complexity" than most people possess in terms of multi-modal sensory input does not result in a less complex conscious experience. That is, consciousness does not appear to be a consequence of an entity possessing a rich sensory window onto the world. Helen Keller shows us that even with extreme deficits in sensory input, consciousness and complex cognition may remain unaffected.

Our second example is that of the many case studies that have been gathered in the Interesting Brains Project (EvLab, MIT). Its goal is to highlight and research people with brains that fall outside of the typical architectures that have informed cognitive neuroscience theory relating architecture and function. We will highlight a single case here as illustrative of the project's findings but encourage anyone to explore the project to get a sense of the exquisite variety of brains living happily in the world today.

The individual EG (in neuropsychology research participants are commonly identified using initials) was born without a left temporal lobe (Tuckute 2022). If you asked a cognitive neuroscientist what types of deficits to expect given this information, the list would run long. It would be a commonsense expectation that EG would have severely compromised language abilities and language understanding, based on the huge body of literature that implicates the specialization of the left cortical hemisphere for language processing, especially left temporal and left prefrontal cortex (Josse and Tzourio-Mazoyer 2004). However, EG wrote a personal introduction to the article cited above describing her case study, stating that she's completed graduate school and has learned Russian as a second language so well that she has dreamed in it.

Many other striking cases like the above can be found, including a 44-year-old tax professional who discovered via an incidental MRI that he only had roughly 5% of a brain in his skull due to hydrocephalus (Schwarcz 2017), and the hydranencephalic children discussed by neuropsychologist Mark Solms (Solms 2013). What we can reasonably conclude from these cases of interesting brains and minds is that consciousness does not appear to be a consequence of a specific pattern of cognitive architecture. Numerous case studies from the Interesting Brains Project and others (e.g., Kofman and Levin 2025) demonstrate that spectacularly different brain architectures nonetheless produce conscious humans capable of complex cognition. AI researchers, who often talk about "the human brain" and "the capabilities of human intelligence" need to know that the mapping between brain hardware and functional cognitive performance, not to mention inner perspective, is not nearly a solved problem in neuroscience. Additionally, data on memories surviving regeneration of the brain and being moved via tissue transplants shed doubt on the simple conventional story of intelligence, which has massive implications for AI (Blackiston, Shomrat, and Levin 2015).

A third and more mundane example is color blindness, which has a worldwide prevalence of roughly 8% in males and 0.5% in females (Birch 2012). Given the reduced sensory complexity of color-blind individuals, would we expect them to be any "less conscious" then the majority of trichromats? Absolutely not. How far can we follow this argument? What if someone loses their sense of smell? What if they lose their arms and legs? Would any of this impact their consciousness? We don't believe so. The implication is that while cognitive complexity is certainly a nice-to-have, consciousness seems to exist to at least some extent independently of it.

Because of widespread tendency to conflate cognitive complexity and consciousness, AI researchers often characterize machine consciousness as something that's "just around the corner" because of how much AI cognitive complexity has increased – in other words, we take it for granted that consciousness is not just associated with complexity but is a *consequence* of complexity. However, looking once again at the case of Helen Keller we might instead say that consciousness is *using* cognition. That is, consciousness itself may be an agentic "force" that leverages whatever cognitive, biological, or inorganic mechanisms are available to it. If this is indeed the relationship between consciousness and cognition, it has strong implications for how we approach the engineering of consciousness in AI systems. We might need to consider ways to "lasso" consciousness where it already exists and perhaps provide it with artificial intelligence mechanisms to use, rather than build the intelligence mechanisms first and expect consciousness to somehow emerge from them. A number of efforts in the bioengineering of hybrots and cyborgs have revealed new technological paths for creating novel bodies to examine the capabilities of natural and synthetic beings (Clawson and Levin 2023).

Behavioral Cognitive Complexity

The third and final assumption we wish to address concerns the detection and measurement of consciousness through the observation of behavior. We focus on the wide range of behavioral cognitive complexity exhibited by conscious humans and its relationship to consciousness. Designing valid and reliable instruments to measure consciousness is challenging, to say the least. We must make inferences that reasonably connect an entity's observable behavior with some hypothesized latent (i.e., itself unobservable) construct of consciousness that we take to be a property of that entity. The modern field of psychometrics is entirely concerned with the problem of latent variable modeling, and if nothing else we recommend a dive into psychometric theory and methodology for anyone who is attempting to design a detection measure for machine consciousness (Furr 2021).

Examples of behavioral complexity, plasticity, and flexibility are ubiquitous. As with architectural complexity, we are prone to work backward from this complexity to argue that it counts as evidence for consciousness, and to design benchmarks for AI systems that require them to produce a semblance of that behavioral complexity. If an AI system could produce language, if an AI system could demonstrate agency, if an AI system could make a good painting…each of these and many other tests have been proposed as benchmarks for AGI, superintelligence, and ultimately consciousness. However, the goal posts keep moving (Hendrycks et al. 2025); Turing tests that were once held up as definitive are now easily passed, and still not everyone is convinced that we are any closer to machine consciousness than before. We believe this is revealing of how behavioral cognitive complexity is implicitly understood by many people to be missing the mark–when these benchmarks are exceeded, some of us keep asking for more, feeling like maybe we didn't get the evaluation criteria quite right yet.

What might help to clarify this point is an example of human behavioral *inflexibility*, or limited range. Consider the Amazon fulfillment center worker. Stowers, pickers, and packers of goods wear wrist- and finger-mounted scanning devices, which transmit instructions to the worker for where to go and what items to select, and operate on countdown timers to encourage workers to meet performance targets–which are also recorded by the scanners and transmitted in real time back to the workers (Gent 2025). This form of algorithmic management results in an extremely limited and stereotyped set of behaviors within the fulfillment center, with little human-to-human communication as the algorithm is optimized for efficiency and prevents people from occupying the same space at the same time. The effectiveness of this process is based on an exploitation of exceedingly complex, agentic human beings that purposefully restricts their degrees of freedom and instruments them for the objective of warehouse efficiency. If we imagined an alien species witnessing this fulfillment center, what conclusions might they come to about the likelihood of consciousness and complex cognition existing in the humans working there?

When we observe in nature a seemingly simple entity exhibiting a seemingly simple behavioral repertoire, how likely are we to consider that a fully-fledged consciousness might have been cornered by evolutionary and environmental pressures into expressing a small sliver of its potential capabilities? More commonly, we interpret it at face value and conclude that the entity is only *capable* of expressing the behavioral repertoire that we observe. We think that if that entity could do more, it would do more, without considering that its intelligence does not exist in a vacuum. Like the Amazon workers, all entities in the universe exist in relation to one another. Sometimes we are constrained in ways that are not under our control, in ways that limit the expression of our fullest potential. An extreme case of this can be found in patients with pseudocoma or "locked-in syndrome", who exhibit complete or near-complete paralysis with otherwise generally intact cognition and consciousness (Smith and Delargy 2005). This condition can be extremely difficult to diagnose, which again points to the challenge we face in disentangling the presence of consciousness with the observability of complex behavior. As with architectural cognitive complexity, we believe the implications of this for machine consciousness are profound, in that we cannot reliably depend on measures of complex behavior to provide us with definitive evidence in support of or against the presence of consciousness.

Away from Complexity

What do we find if we turn towards simplicity? What about the cognitive "basics", so to speak? An accumulating body of evidence in neuroscience points to structures in the human brain that *do* seem to have a causal relation to conscious awareness. These structures do not support cognitive processes that are traditionally thought of as complex, however, like language, deductive reasoning, or imagination. Instead, this research points to structures in the brainstem that provide us with internal feedback on the state of the body and support the rather underappreciated sense of *interoception* (Craig 2003). It is worth noting that while Helen Keller and the other people with interesting brains mentioned here may have radically different neural architectures and/or reduced access to typical sensory input, we are not aware of any cases in which a person has lived to adulthood without a brainstem. Thus, an alternative conclusion from their stories is that consciousness persists in all these people because of their intact brainstems and intact interoception. This view is essentially identical with the *affective-interoceptive* theory of consciousness, developed by neuroscientist Antonio Damasio and based on evidence for a causal relationship between basic awareness and somatic maps of the state of the body in the brainstem (Damasio & Damasio 2023). This theory points us away from thinking of consciousness as arising from cognitive complexity, and towards physical embodiment. Not surprisingly, this point of view resonates particularly with the robotics community, and artificial interoceptive/proprioceptive modeling is an active area of AI research there (Bongard, Zykov, and Lipson 2006; Chen et al. 2022).

Additional evidence for the abilities of relatively simple systems can be found in the emerging field of Diverse Intelligence, in which learning and problem-solving competencies have been demonstrated in all kinds of non-neural biologicals (Baluška and Levin 2016), inorganic materials (Kaygisiz and Ulijn 2025), and other systems. These data rarely inform current debates on AI, much like the facts of developmental and evolutionary biology which emphasize the need for models of scaling mind from its chemical origins–as Turing (1953) well understood–not for binary categories of "is" or "isn't" conscious. Moreover, consideration of our own evolutionary and developmental history, as well as increasing recognition of problem-solving capacities performed by cells and tissues in physiological and other spaces, reminds us that inability to make verbal reports using language does not reliably rule out consciousness. The proposed existence of consciousness in relatively simple systems, and the continuity of developmental, evolutionary, and engineering processes suggests that consciousness may be a graded property. If it is, it further complicates some of the examples from the human case studies presented above in that we may need to consider different people possessing different degrees of consciousness, or even different degrees of consciousness existing within the same individual in different states or periods of life and development.

The examples we have shared challenge the field of AI research to revisit its motivations to increase the cognitive complexity of AI systems with the goal of producing machine consciousness. For instance, there are recent efforts to implement multimodal sensor-augmented large language models to facilitate "context aware" functioning (Li et al. 2025). It is tempting to envision that this is a step towards machine consciousness, and maybe it is. We are concerned that it may instead be a kind of off-ramp, and that such efforts may be locked in a feedback cycle seeking to increase architectural and functional cognitive complexity in AI systems with no certain end in sight. What if we can add attention? (de Santana Correia and Colombini 2022). Emotion? (Assuncao et al. 2022). Meta-cognition? (Johnson et al. 2024). Motivation? (O'Reilly 2020). The list of potential cognitive targets is long, long enough to keep us busy for decades without tackling the problem of consciousness head-on. Instead, we support leveraging neural and non-neural biology to understand better the mapping between consciousness and an observed material interface.

Conclusion

We encourage AI researchers to deeply consider their underlying arguments for adding cognitive complexity to AI systems and to question whether that will truly lead us where we want to go when consciousness is the goal. Nature has been deploying intelligence long before neurons came on the scene, and a mature theory of consciousness and its embodiments is surely going to have to say something about its kind and degree in a very wide range of natural, artificial, and hybrid systems (starting below the single-cell level).

We emphasize the fundamentally tenuous relationship between architectural or behavioral complexity and consciousness; however, we acknowledge that our own human experience of complex cognition might have an influence on our ability to recognize or interact with machine consciousness in practicality. That is, even if a machine consciousness were hypothetically proven to exist in a very simple system, we as humans with complex cognitive architecture and behavior might have difficulty in relating to it. Perhaps our instincts towards developing more cognitively complex AI reveal our desires to interact with a consciousness that is more recognizably like our own. It is reasonable to think that there must be some kind of "impedance match" or resonance between a mind and the kind of things it can observe and recognize; if true however, it only reinforces the need to develop tools (conceptual, and empirical) to augment our native limitations to be able to detect, and ethically relate to, a wider range of minds.

In reviewing some key assumptions that pervade our thinking on consciousness and cognitive complexity, we invite new kinds of discussion among AI researchers. To see consciousness research growing within the field of AI is exciting, and we look forward to participating in the creative exploration of a new future where conscious humans and machines may co-exist to the benefit of all.

I like to think (and
the sooner the better!)
of a cybernetic meadow
where mammals and computers
live together in mutually
programming harmony
like pure water
touching clear sky.

Excerpt from the poem
All Watched Over by Machines of Loving Grace
by Richard Brautigan

Acknowledgements

We would like to thank David Burke, who shared valuable insights over multiple conversations on the quantum mechanics frameworks discussed, and Raj Patra for his expert pedagogy and engaging discussions on Tantric philosophy.

The authors attest that no AI tools were used to draft this manuscript.

References

Arbib, M. A. 2001. Co-evolution of Human Consciousness and Language. *Annals of the New York Academy of Sciences* 929(1): 195–220. doi.org/10.1111/j.1749-6632.2001.tb05717.x.

Assunção, G.; Patrão, B.; Castelo-Branco, M.; and Menezes, P. 2022. An Overview of Emotion in Artificial Intelligence. *IEEE Transactions on Artificial Intelligence* 3(6): 867–886. doi.org/10.1109/tai.2022.3159614.

Baars, B. J., and McGovern, K. 2003. Cognitive Views of Consciousness: What are the Facts? How Can we Explain Them? In *The Science of Consciousness: Psychological Neuropsychological and Clinical Reviews*, edited by M. Velmans, 75–107. London: Routledge.

Baluška, F., and Levin, M. 2016. On Having No Head: Cognition Throughout Biological Systems. *Frontiers in Psychology* 7: 902. doi.org/10.3389/fpsyg.2016.00902.

Behrendt, R. P. 2013. Conscious Experience and Episodic Memory: Hippocampus at the Crossroads. *Frontiers in Psychology* 4: 304. doi.org/10.3389/fpsyg.2013.00304.

Birch, J. 2012. Worldwide Prevalence of Red-Green Color Deficiency. *Journal of the Optical Society of America A* 29(3): 313–320. doi.org/10.1364/josaa.29.000313.

Blackiston, D. J.; Shomrat, T.; and Levin, M. 2015. The Stability of Memories During Brain Remodeling: A Perspective. *Communicative & Integrative Biology* 8(5): e1073424. doi.org/10.1080/19420889.2015.1073424.

Chen, B.; Kwiatkowski, R.; Vondrick, C.; and Lipson, H. 2022. Fully Body Visual Self-modeling of Robot Morphologies. *Science Robotics* 7(68): eabn1944. doi.org/10.1126/scirobotics.abn1944.

Clawson, W. P., and Levin, M. 2023. Endless Forms Most Beautiful 2.0: Teleonomy and the Bioengineering of Chimaeric and Synthetic Organisms. *Biological Journal of the Linnean Society* 139(4): 457–486. doi.org/10.1093/biolinnean/blac073.

Craig, A. D. 2003. Interoception: The Sense of the Physiological Condition of the Body. *Current Opinion in Neurobiology*, 13(4): 500–505. doi.org/10.1016/s0959-4388(03)00090-4.

Damasio, A., and Damasio, H. 2023. Feelings Are the Source of Consciousness. *Neural Computation* 35(3): 277–286. doi.org/10.1162/neco_a_01521.

de Santana Correia, A., and Colombini, E. L. 2022. Attention, Please! A Survey of Neural Attention Models in Deep Learning. *Artificial Intelligence Review* 55(8): 6037–6124. doi.org/10.1007/s10462-022-10148-x.

EvLab. (n.d.). *Interesting Brains Project*. Massachusetts Institute of Technology. https://www.evlab.mit.edu/interesting-brains-project. Accessed 2026-01-26.

Fuchs, C. A.; Mermin, N. D.; and Schack, R. 2014. An Introduction to QBism with an Application to the Locality of Quantum Mechanics. *American Journal of Physics* 82(8): 749–754. doi.org/10.1119/1.4874855.

Furr, R. M. 2021. *Psychometrics: An Introduction*. Los Angeles: Sage.

Gent, C. 2025. Under the Scan Gun: Algorithmic Worker Management in English Coal Country. *LOGIC(S) magazine*, 23, 77–82.

Hendrycks, D.; Song, D.; Szegedy, C.; Lee, H.; Gal, Y.; Brynjolfsson, E.; Li, S. ... and Bengio, Y. 2025. A Definition of AGI. arXiv preprint. arXiv:2510.18212 [cs.AI]. doi.org/10.48550/arXiv.2510.18212.

Hoffman, D. 2019. *The Case Against Reality: Why Evolution Hid the Truth from Our Eyes*. New York: WW Norton & Company.

Johnson, S. G.; Karimi, A. H.; Bengio, Y.; Chater, N.; Gerstenberg, T.; Larson, K.; Levine, S. ... and Grossmann, I. 2024. Imagining and Building Wise Machines: The Centrality of AI Metacognition. arXiv preprint. arXiv:2411.02478 [cs.AI]. doi.org/10.48550/arXiv.2411.02478.

Josse, G., and Tzourio-Mazoyer, N. 2004. Hemispheric Specialization for Language. *Brain Research Reviews* 44(1): 1–12. doi.org/10.1016/j.brainresrev.2003.10.001.

Kaygisiz, K., and Ulijn, R. V. 2025. Can Molecular Systems Learn?. *ChemSystemsChem* 7(2): e202400075. doi.org/10.1002/syst.202400075.

Keller, H. 2000. *The Story of My Life.* Illinois: Project Gutenberg. https://www.gutenberg.org/ebooks/2397. Accessed 2026-01-26.

Kofman, K., and Levin, M. 2025. Cases of Unconventional Information Flow Across the Mind-Body Interface. *Mind and Matter* 23(1): 13–69. doi.org/10.53765/mm2025.13.

Levin, M. 2025a. Ingressing Minds: Causal Patterns Beyond Genetics and Environment in Natural, Synthetic, and Hybrid Embodiments. PsyArXiv preprint. PsyArXiv: 5g2xj_v3. doi.org/10.31234/osf.io/5g2xj_v3.

Levin, M. 2025b. The Platonic Space Symposium. https://thoughtforms.life/symposium-on-the-platonic-space/. Accessed 2026-01-27.

Li, Z.; Deldari, S.; Chen, L.; Xue, H.; and Salim, F. D. 2025. Sensorllm: Aligning Large Language Models with Motion Sensors for Human Activity Recognition. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: 354–379. doi.org/10.18653/v1/2025.emnlp-main.19.

Lovejoy, A. 1936. *The Great Chain of Being: A Study of the History of an Idea*. Cambridge: Harvard University Press.

Merker, B. 2007. Consciousness Without a Cerebral cortex: A Challenge for Neuroscience and Medicine. *Behavioral and Brain Sciences* 30(1): 63–81. doi.org/10.1017/s0140525x07000891.

Michel, M. 2022. Conscious Perception and the Prefrontal Cortex: A Review. *Journal of Consciousness Studies* 29(7): 115–157. https://doi.org/10.53765/20512201.29.7.115.

Neuman, Y., and Nave, O. 2010. Why the Brain Needs Language in Order to be Self-Conscious. *New Ideas in Psychology* 28(1): 37–48. doi.org/10.1016/j.newideapsych.2009.05.001.

O'Reilly, R. C. 2020. Unraveling the Mysteries of Motivation. *Trends in Cognitive Sciences* 24(6): 425–434. doi.org/10.1016/j.tics.2020.03.001.

Penfield, W. 1938. The Cerebral Cortex in Man: I. The Cerebral Cortex and Consciousness. *Archives of Neurology & Psychiatry* 40(3): 417–442. doi.org/10.1001/archneurpsyc.1938.02270090011001.

Planck, M. 1931, January 25. *The Observer*, London, England. https://en.wikiquote.org/wiki/Max_Planck. Accessed 2026-01-27.

Rovelli, C. 1996. Relational Quantum Mechanics. *International Journal of Theoretical Physics* 35(8): 1637–1678. doi.org/10.1007/bf02302261.

Schlegel-O'Brien, K. 2024. Materialism Matters: The Role of Philosophy in Science. Advanced Science News. https://www.advancedsciencenews.com/materialism-matters-the-role-of-philosophy-in-science/. Accessed 2026-03-03.

Schwarcz, J. 2017. Can You Live Without a Brain? McGill University. https://www.mcgill.ca/oss/article/health-you-asked/it-true-you-can-live-without-brain. Accessed 2026-01-26.

Searle, J. R. 2002. *Consciousness and Language*. Cambridge: Cambridge University Press.

Shields, C. J. 2016. *Aristotle: De Anima*. Oxford: Oxford University Press.

Smith, E., and Delargy, M. 2005. Locked-in Syndrome. *BMJ* 330(7488): 406–409. doi.org/10.1136/bmj.330.7488.406.

Solms, M. 2013. The Conscious Id. *Neuropsychoanalysis* 15(1): 5–19. doi.org/10.1080/15294145.2013.10773711.

Tuckute, G.; Paunov, A.; Kean, H.; Small, H.; Mineroff, Z.; Blank, I.; and Fedorenko, E. 2022. Frontal Language Areas do not Emerge in the Absence of Temporal Language Areas: A Case Study of an Individual Born Without a Left Temporal Lobe. *Neuropsychologia* 169: 108184. doi.org/10.1101/2021.05.28.446230.

Turing, A. M. 1990. The Chemical Basis of Morphogenesis 1953. *Bulletin of Mathematical Biology* 52(1): 153–197. doi.org/10.1016/s0092-8240(05)80008-4.

Watson, R. A.; Levin, M.; and Buckley, C. L. 2022. Design for an Individual: Connectionist Approaches to the Evolutionary Transitions in Individuality. *Frontiers in Ecology and Evolution* 10: 823588. https://doi.org/10.3389/fevo.2022.823588.